

# Generalized Maximum Entropy Estimation of Discrete Sequential Move Games of Perfect Information\*

Yafeng Wang and Brett Graham<sup>†</sup>

Wang Yanan Institute for Studies in Economics (WISE)

Xiamen University

December 7, 2010

## Abstract

We propose a data-constrained generalized maximum entropy estimator for discrete sequential move games of perfect information. Unlike most other work on the estimation of complete information games, the method we proposed is data constrained and requires no simulation or assumptions about the distribution of random preference shocks. We formulate the GME estimation as a (convex) mixed-integer nonlinear optimization problem which can be easily implemented on optimization software with high-level interfaces such as GAMS. The model is identified with only weak scale and location normalizations. Monte Carlo evidence demonstrates that the estimator can perform well in moderately size samples. As an application we study the location choice of German siblings using the German Ageing Survey.

**Keywords:** Game-Theoretic Econometric Models, Sequential-Move Game, Generalized Maximum Entropy, Mixed-Integer Nonlinear Programming.

**JEL Classification Numbers:** C01, C13, C35, C51, C72.

---

\*First version: Dec. 2009.

<sup>†</sup>Yafeng Wang: [econwang@econwang.co.cc](mailto:econwang@econwang.co.cc). Brett Graham: [bgraham@xmu.edu.cn](mailto:bgraham@xmu.edu.cn).

# 1 Introduction

This paper focuses on the estimation of discrete sequential-move games of complete information. The estimation of complete information sequential-move games have received less attention than their incomplete information counterparts <sup>1</sup>. since they involve multidimensional integrals which incorporates a higher degree of computational complexity. Recently, [Bajari, Hong, and Ryan \(2010\)](#) have provided a simulation-based estimator for static complete information discrete games based on importance sampling. Estimation of the general class of sequential-move (extensive form) games has been quite limited <sup>2</sup>.

To the best of our knowledge, [Maruyama \(2009\)](#) is the only existing methodology for estimating such games. This estimator is a modified version of the GHK simulator (Geweke (1989, 1991), Hajivassiliou and McFadden (1998) and Keane (1990, 1994)) and is essentially a maximum simulated likelihood (MSL) estimator. As is well known, MSL is biased for any fixed number of simulations (e.g. [Lee, 1995](#)). In order to obtain  $\sqrt{T}$  consistent estimators, one needs to increase the number of draws  $S$  so that  $\frac{S}{\sqrt{T}} \rightarrow \infty$ . Such an estimator requires large scale simulation making its computational burden high. It also requires that the random preference shocks are normally distributed making its scope for application limited.

We overcome these computational and theoretical constraints by using the generalized maximum entropy (GME) approach<sup>3</sup>.

Our data-constrained GME estimator is obtained by maximizing the entropy-information function of Shannon (1948) subject to the model property constraints (such as equilibrium conditions). Due to the structure of perfect information sequential-move

---

<sup>1</sup>Studies of incomplete information simultaneous-move games (normal form) and dynamic games include ([Bajari, Hong, Krainer, and Nekipelov, 2010](#); [Aguirregabiria and Mira, 2007](#))

<sup>2</sup>[Berry \(1992\)](#), [Mazzeo \(2002\)](#) and [Schmidt-Dengler \(2006\)](#) estimate some simplified sequential-move games with special game structures.

<sup>3</sup>The GME principle was introduced by [Golan, Judge, and Miller \(1996\)](#), which is based on the classic maximum entropy (ME) approach of [Jaynes \(1957a,b\)](#). Further econometric studies using the GME approach include ([Golan, Judge, and Miller, 1996](#); [Mittelhammer and Cardell, 1997](#); [Golan, Perloff, and Shen, 2000](#); [Campbell and Hill, 2001](#)).

game, the equilibrium conditions (specifically sub-game perfection) contain logical connections between the endogenous variables. The resulting constrained optimization problem can be viewed as a mixed-integer nonlinear optimization problem, since the logical statements can be reinterpreted as binary variables (Williams, 1985). The problem can then be solved with an efficient algorithm such as Branch and Bound (BB), Outer-Approximation (OA) and or OA based Branch-and-Cut (B-Hyb). As shown below, with a linear payoff function, our GME problem becomes a convex MINLP, which can be exactly solved by many existing algorithms (Bonami, Kilinc, and Linderoth 2009).

Our approach makes several contributions to the literature on estimating game theoretic models of complete (perfect) information. First, by using the data constraints instead of moment constraints we avoid the problems associated with multidimensional integrals which makes the computational burden of our approach acceptable for most applications. Although we focus on the sequential-move game, our approach can be extended to static game of complete information<sup>4</sup>. Second, our approach makes no assumption about the distributional form of random preference shocks. In contrast, existing estimators for general complete information games, such as Bajari, Hong, and Ryan (2010) for the static case and Maruyama (2009) for the sequential-move case require more restrictive assumptions. Bajari, Hong, and Ryan (2010) requires that the form of the distribution function should be known. Maruyama (2009) requires that the distribution function should be normal. Third, we reformulate the estimation problem as a MINLP since there are logical connections between the endogenous variables under the equilibrium conditions. To the best of our knowledge, our estimator is the first one to make use of MINLP in econometric estimation problems<sup>5</sup>. Our Monte Carlo results show the validity of this estimation procedure, which can be extended to other estimation problem where logical statements

---

<sup>4</sup>Golan, Karp, and Perloff (1998, 2000) also make use of the GME to estimate the static game. Their constraints are moment based since they deal with the incomplete information case.

<sup>5</sup>Jouneau-Sion and Torres (2006) formulate the Maximized Monte Carlo (MMC) test as a Mixed Integer Programming problem.

arise.

The greatest shortcoming of our approach is that it is difficult if not impossible to construct the large sample properties of our estimators. As argued by [Su and Judd \(2008\)](#), in order to compute standard errors, we need to formulate the exact Hessian matrix of the objective function with respect to the structural parameters. Such work seems difficult within the MINLP framework. However, since our Monte Carlo simulations show consistency and asymptotic normality of the proposed estimator, we use the paired bootstrap method to construct standard errors. Following the arguments of [Horowitz \(1997, 1998, 2001\)](#), [Campbell and Hill \(2001\)](#) and [Su and Judd \(2008\)](#), this method may consistently estimate the distribution of our estimator.

The paper is organized as follows. In section 2 we outline the general discrete sequential-move game to be estimated and formulate its equilibrium conditions. A simple  $2 \times 2 \times 2$  sequential entry game which will be used for expositional purposes is described. A brief review of maximum entropy, generalized maximum entropy estimation and (convex) mixed-integer nonlinear programming are presented in section 3. We discuss the identification issue from the nature of the game structure and equilibrium conditions in section 4. Our GME estimation for the discrete sequential-move game of perfect information is also presented. Monte Carlo simulations are conducted in section 5. Section 6 concludes the paper.

## 2 The Model

There are  $T$  independent repetitions of a sequential-move game of perfect information. In each game there are  $i = 1, \dots, N$  players. Since we study sequential-move games of perfect information we assume that player  $i$  chooses  $a_i$  from the finite set of actions  $A_i$  after observing  $i - 1$  actions. Define  $A = \times_i A_i$  and let  $a = (a_1, \dots, a_N)$  denote a generic element of  $A$ . Player  $i$ 's von Neumann-Morgenstern (vNM) utility is a mapping  $u_i : A^N \rightarrow R$ ,

where  $R$  is the real line. We will sometimes drop the subscript  $t$  for simplicity when no ambiguity would arise.

Following [Bresnahan and Reiss \(1990, 1991\)](#), assume that the vNM utility of player  $i$  can be written as:

$$u_i(a, x, \epsilon_i; \theta) = f_i(x, a; \theta) + \epsilon_i(a) \quad (1)$$

where  $a \in A^N$ . In Equation (1), player  $i$ 's vNM utility from outcome  $a$  is the sum of two terms. The first term  $f_i(x, a; \theta)$  is a function which depends on outcome  $a$ , the vector of actions taken by all of the players,  $x$ , the players' characteristics and some other variables which influence utility, and parameters  $\theta$ , covariates  $x$  are observed by the econometrician. The second term  $\epsilon_i(a)$ , is a random preference shock which reflects the information about utility that is common knowledge to the players but not observed by the econometrician. Unlike [Maruyama \(2009\)](#), here the preference shocks depend on the entire vector of actions  $a$ , not just the actions taken by player  $i$ . As argued by [Bajari, Hong, and Ryan \(2010\)](#), this is a more general setting. For example, in a simple entry game, the unobserved information about one player's payoff to the econometrician may not only vary across his own actions but also across the actions of other players.  $\epsilon_i(a)$  are assumed to be independent or have some known dependence. Let  $\epsilon_i$  denote the vector of the individual shocks  $\epsilon_i(a)$  and  $\epsilon$  denote the vector of all shocks. We will discuss more about the structure of  $\epsilon_i$  in the identification and estimation section.

A strategy for player  $i$  is a mapping from  $A^{i-1}$  to  $A_i$ , where  $A^i = \times_{j \leq i} A_j$ . The equilibrium concept corresponding to sequential-move games of perfect information, subgame perfection (SPE), is a strategy profile in which every player expects no gain from individually deviating from their equilibrium in every subgame.

Formally, an action profile,  $a^{SPE} = (a_1^{SPE}, \dots, a_i^{SPE}, \dots, a_N^{SPE})$ , is an SPE if

$$u_i(a^{SPE}, x, \epsilon_i; \theta) - u_i(a_{<i}^{SPE}, a_i, a_{>i}^*(a_{<i}^{SPE}, a_i), x, \epsilon_i; \theta) \geq 0 \quad (2)$$

*for all  $i = 1, \dots, N$  and all  $a_i \neq a_i^{SPE}$ ,*

where  $a_{>i}^*(a_{\leq i})$  is an action profile resulting from the subgame that starts from player  $i + 1$  given the decisions of the preceding players,  $a_{\leq i}$  and where each player in the subgame uses the same strategy that generated the action profile  $a^{SPE}$ . These equilibrium conditions are defined recursively and the solution can be easily calculated by the backward induction for any given set of parameters  $\theta$ , observed covariates,  $x$ , and unobserved shocks  $\epsilon$ . Kuhn's theorem ensures the existence to solutions of the inequality system (2), and thus ensures that every finite sequential move game of perfect information has an SPE. As noted by [Berry and Tamer \(2007\)](#), dealing with multiple equilibria complicates the identification problem. Fortunately, every finite sequential-move game with perfect information in which no player is indifferent between any two outcomes has a unique subgame perfect equilibrium ([Osborne and Rubinstein \(1994\)](#)). Since  $\epsilon_i(a)$  has an atomless distribution in our econometric model, we can ignore the indifference case.

Given this structure of the discrete choice sequential-move game, our task is to estimate and draw an inference about the parameters of the payoff functions,  $\theta$ , given a set of observed action profiles  $a^o$ , some covariates which effect the payoffs,  $x$ , and an exogenous decision order. Note that the actual payoff levels are latent variables.

For the purpose of exposition, we provide a simple sequential entry game as an example. There are two players who act as potential entrants in each of  $T$  markets. The structure of this entry game is illustrated in figure [a](#) with payoffs  $u_1$  and  $u_2$  as defined in Equation (1).  $a_i \in \{0, 1\}$  is firm  $i$ 's action, 0 means no entry and 1 means entry. For  $(0, 0)$

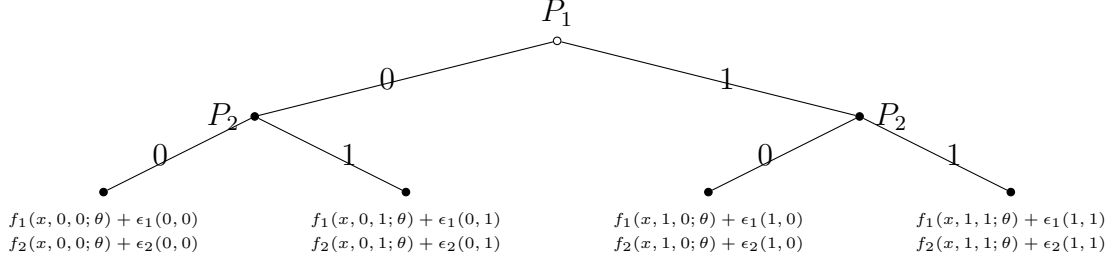


Figure a: A simple entry game

to be an equilibrium outcome, we require the following three equilibrium conditions:

$$\begin{aligned}
u_1(x, 0, 0, \epsilon_1(0, 0); \theta) &> u_1(x, 1, 0, \epsilon_1(1, 0); \theta) \quad \text{if} \quad u_2(x, 1, 0, \epsilon_2(1, 0); \theta) > u_2(x, 1, 1, \epsilon_2(1, 1); \theta) \\
u_1(x, 0, 0, \epsilon_1(0, 0); \theta) &> u_1(x, 1, 1, \epsilon_1(1, 1); \theta) \quad \text{if} \quad u_2(x, 1, 0, \epsilon_2(1, 0); \theta) \leq u_2(x, 1, 1, \epsilon_2(1, 1); \theta) \\
u_2(x, 0, 0, \epsilon_2(0, 0); \theta) &\leq u_2(x, 0, 1, \epsilon_2(0, 1); \theta)
\end{aligned} \tag{3}$$

The equilibrium conditions required for the other three action profiles to be equilibrium outcomes can be formulated similarly. Since the econometrician only observes the equilibrium outcomes but not strategies, the equilibrium conditions can be viewed as logical statements that are conditional on the off equilibrium path choices, which are unobserved by the econometrician. We make use of MINLP to handle such logical statements.

### 3 Preliminary

We start by providing some background to the GME approach. Also, since our GME estimator is obtained via a (convex) MINLP, we also provide a review of the MINLP problem.

#### 3.1 A Basic Review of GME

In the classic maximum entropy (ME) approach, [Shannon \(1948\)](#)'s entropy is used to

measure the uncertainty (state of knowledge) we have about the occurrence of a collection of events. Let  $x$  be a random variable with possible outcomes  $x_s$ ,  $s = 1, 2, \dots, n$ , with probabilities  $\alpha_s$  such that  $\sum_s \alpha_s = 1$ . [Shannon \(1948\)](#) defined the entropy of the distribution  $\alpha = (\alpha_1, \dots, \alpha_n)'$ , as

$$H \equiv - \sum_s \alpha_s \ln \alpha_s \quad (4)$$

where  $0 \ln 0 \equiv 0$ . To recover the unknown probabilities  $\alpha$  that characterize a given data set, [Jaynes \(1957a,b\)](#) proposed maximizing entropy, subject to the available sample-moment information and adding up constraints on the probabilities.

Within the classic ME framework, the observed moments are assumed to be exact. To extend this approach to the problem with noise, the GME approach (developed by [Golan, Judge, and Miller \(1996\)](#)) generalizes the ME approach by using a dual objective (precision and prediction) function. We illustrate the GME approach via a linear model:

$$Y = X\beta + \varepsilon \quad (5)$$

where  $Y$  is an  $N \times 1$  dependent variable vector,  $X$  is an  $N \times K$  matrix of explanatory variables,  $\beta$  is a  $K \times 1$  a vector of parameters, and  $\varepsilon$  is an  $N \times 1$  vector of disturbance terms. The GME estimator of  $\beta$  in this general linear model is derived in the following manner. First,  $\hat{p} = (\hat{p}'_1, \dots, \hat{p}'_K)'$  is derived from the following constrained maximum entropy



problem:

$$\begin{aligned}
& \max_{p'_k, w'_i} - \sum_{k=1}^K p'_k \ln(p_k) - \sum_{i=1}^N \omega'_i \ln(\omega_i) \\
& s.t. \quad Y = XZp + V\omega \\
& \quad 1'p_k = 1, \quad \forall k \\
& \quad 1'\omega_i = 1, \quad \forall i \\
& \quad p_k > [0], \quad \omega_i > 0. \quad \forall i, k.
\end{aligned} \tag{6}$$

$Z$  and  $V$  are  $K \times KM$  and  $N \times NJ$  matrices of support points for the  $\beta$  and  $\varepsilon$  vectors, defined respectively, as:

$$Z = \begin{bmatrix} z'_1 & 0 & \dots & 0 \\ 0 & z'_2 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & z'_K \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} v'_1 & 0 & \dots & 0 \\ 0 & v'_2 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \cdot & v'_N \end{bmatrix} \tag{7}$$

The  $M \times 1$  vector  $z_k = (z_{k1}, \dots, z_{kM})'$  is such that  $z_{k1} < z_{k2} \leq \dots \leq z_{kM}$  and it is assumed that  $\beta_k \in (z_{k1}, z_{kM})$  for each  $k$ <sup>6</sup>. The vector  $v_i = (v_{i1}, \dots, v_{iJ})'$  is such that  $v_{i1} < v_{i2} \leq \dots \leq v_{iJ}$  and it is assumed that  $\varepsilon_i \in (v_{i1}, v_{iJ})$  for each  $i$ . Typically, each  $v_{i1}$  and  $v_{iJ}$  will be uniformly and symmetrically distributed about zero and have the same  $J$  dimensions. The actual bounds used for a given problem depend on the observed sample as well as any available conceptual or empirical information<sup>7</sup>.  $p_k = (p_{k1}, \dots, p_{kM})'$  and

---

<sup>6</sup>This parameter support is based on prior information or economic theory, for example, we might specify boundaries of  $z_{k1} = 0$  and  $z_{kM} = 1$  when estimating the marginal propensity to consume. Without any available prior information, we can specify  $z_k$  to be symmetric around zero, with large negative and positive boundaries, for example,  $z_{k1} = -z_{kM} = -10^6$ .

<sup>7</sup>One viable approach is to use Chebychev's Inequality or the three-sigma rule (Pukelsheim, 1994) and assume the errors are drawn from a uniform distribution with mean zero and variance  $(y_{\max} - y_{\min})/12$ .

$\omega_i = (\omega_{i1}, \dots, \omega_{iJ})'$  are non-negative weight vectors that sum to unity. The GME estimator is then given by  $\hat{\beta} = Z\hat{p}$ . [Golan, Judge, and Miller \(1996\)](#) has a rigorous discussion of this approach and applies it to a rich scopes of econometric problems, such as dynamic models, model selection and discrete choice-censored problems. [Mittelhammer and Cardell \(1997\)](#) establish consistency and asymptotic normality results for the GME estimator under general regularity conditions.

### 3.2 A Basic Review of MINLP

Our GME estimator for the sequential-move game is essentially obtained by solving a generalized disjunctive programming problem, which can be reformulated as an MINLP problem. Here we provide a basic review of the general structure and of existing available algorithms for solving MINLP problems. A simple algebraic representation of an MINLP problem is as follows:

$$\begin{aligned} \min_{\{x,y\}} \quad & Z = f(x, y) \\ \text{s.t.} \quad & g_j(x, y) \leq 0, \quad j \in J \\ & x \in X, y \in Y \end{aligned} \tag{8}$$

where  $f(\cdot)$  and  $g(\cdot)$  are differentiable functions,  $J$  is the index set of inequalities, and  $x$  and  $y$  are continuous and discrete variables, respectively. In most applications the discrete set  $Y$  is restricted to be binary. When both  $f(\cdot)$  and  $g(\cdot)$  are both convex functions, this becomes a convex MINLP problem, which can be exactly solved via most existing algorithms<sup>8</sup>. We take the technique of solving an MINLP problem as given.

---

([Golan, Judge, and Perloff 1997](#))

<sup>8</sup>Methods for solving convex MINLP problems include the branch and bound method (BB) ([Gupta and Ravindran, 1985](#); [Nabar and Schrage, 1991](#); [Borchers and Mitchell, 1994](#); [Stubbs and Mehrotra, 1999](#); [Leyffer, 2001](#)), Generalized Benders Decomposition (GBD) ([Geoffrion, 1972](#)), Outer-Approximation (OA) ([Duran and Grossmann, 1986](#); [Yuan et al., 1988](#); [Fletcher and Leyffer, 1994](#)), LP/NLP based branch and bound ([Quesada and Grossmann, 1992](#)), and Extended Cutting Plane Method (ECP) ([Westerlund and](#)

## 4 Estimation

In this section we first detail the assumptions required for our estimation method. We then describe the estimator in detail before discussing its limitations and extensions.

### 4.1 Assumptions

We require three assumptions in order to use our estimation method. The first is associated with the structure of the game tree, the second is associated with the ordinality of utility, and the third is associated with the distributional properties of the errors terms.

**Assumption 1 (Exogenous Decision Order)** *The decision order of agents in the sequential-move game is exogenous and known to the econometrician.*

Although the exact decision order of agents is rarely observed, to estimate sequential-move games we must assume a specific decision order.

**Assumption 2 (Scale and Location Normalizations)** *The payoff of one action for each player are fixed at a known constant.*

From the equilibrium condition (2), any affine transformation of the payoffs does not perturb the set of equilibria, so scale and location normalizations are necessary. The scale normalization is subsumed in the following assumption about the distribution of the error terms. [Bajari, Hong, and Ryan \(2010\)](#) argue that this restriction is similar to the argument that we can normalize the mean utility from the outside good equal to a constant, usually zero, in a standard discrete choice model.

---

Pettersson, 1995). Methods for solving nonconvex MINLP problems include LP relaxation (Sherali & Adams, 1990), LP and SDP relaxations (Lovasz & Schrijver, 1991), SDP relaxations (Lasserre, 2001) and Branch-and-Reduce (Tawarmalani & Sahinidis, 2002). Such methods are available in some optimization software with high-level interfaces such as GAMS, AMPL, and TOMLAB which has a MATLAB interface. In GAMS, both BARON and BONMIN can solve convex MINLP problems, but only BARON can solve nonconvex MINLP problems.

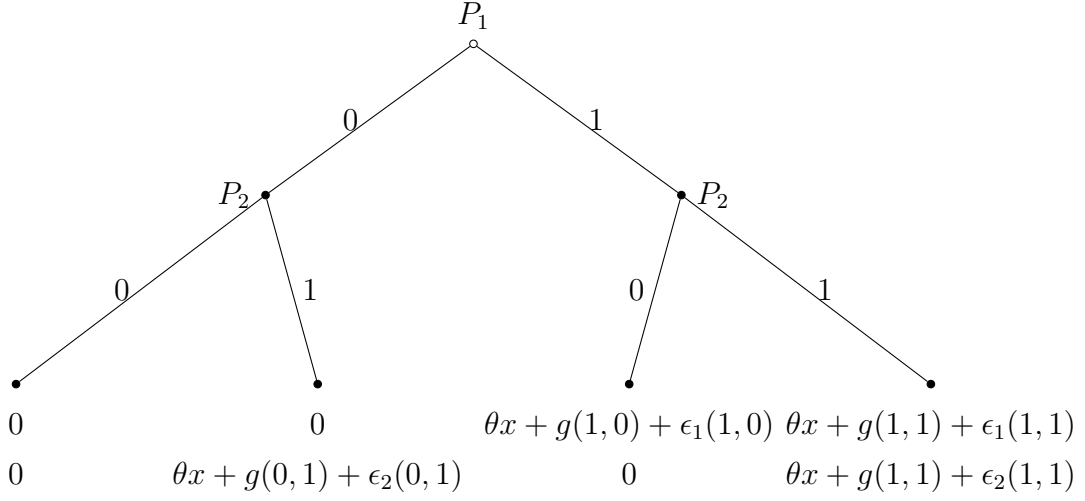


Figure b: A Reformulated Entry Game

**Assumption 3 (Regularity Conditions of Random Shocks)** *The random preference shocks  $\epsilon_{it}(a)$  are i.i.d, are independent of state variables, and have zero mean and finite variance, i.e.  $E(\epsilon_{it}(a)) = 0$ ,  $E(x\epsilon_{it}(a)) = 0$ , and  $Var(\epsilon_{it}(a)) < \infty$ .*

The *i.i.d* assumption is stronger than is required for our GME estimation. Any known heteroskedasticity and dependence among random shocks can be handled within the our framework. We will discuss this further after presenting our estimator. Our GME estimator is semiparametric in the sense that it does not impose any parametric assumption on the random shocks.

For the purpose of exposition, we apply GME estimation under assumption 1 to 3 to the simple entry game which was introduced in section 2. We assume that utility takes the form:

$$u_i(x, a, \epsilon_i; \theta) = 1(a_i = 1)\{\theta x + \delta g(a) + \epsilon_i(a)\} \quad (9)$$

The entry game is shown in figure [b](#).

The location normalization sets the utility from out of the market to 0 and scale normalization sets the parameter  $\delta$  to 1<sup>9</sup>. The equilibrium conditions for action profile  $(0, 0)$  to be an SPE outcomes are:

$$\begin{aligned}
\text{Player1: } & 0 > \theta x + g(1, 0) + \epsilon_1(1, 0) \quad \text{if} \quad 0 > \theta x + g(1, 1) + \epsilon_2(1, 1) \\
& 0 > \theta x + g(1, 1) + \epsilon_1(1, 1) \quad \text{if} \quad 0 \leq \theta x + g(1, 1) + \epsilon_2(1, 1) \\
\text{Player2: } & 0 > \theta x + g(0, 1) + \epsilon_2(0, 1)
\end{aligned} \tag{10}$$

The equilibrium conditions required for the other outcomes can be similarly derived.

## 4.2 The Estimator

In order to use the GME framework, we need to specify the support space  $Z$  for  $\theta$  and the support space  $v_1, v_2, v_3, v_4$  for  $\epsilon_1(1, 0)$ ,  $\epsilon_1(1, 1)$ ,  $\epsilon_2(0, 1)$  and  $\epsilon_2(1, 1)$  respectively. Without loss of generality, all are  $M \times 1$  vectors and  $v_1 = v_2 = v_3 = v_4 = v$ . The corresponding probabilities are defined as  $p^\theta, \omega_t^1, \omega_t^2, \omega_t^3, \omega_t^4$  such that:

$$\theta = \sum_{m=1}^M p_m^\theta z_m \tag{11}$$

$$\epsilon_{1t}(1, 0) = \sum_{m=1}^M \omega_{tm}^1 v, \epsilon_{1t}(1, 1) = \sum_{m=1}^M \omega_{tm}^2 v, \epsilon_{2t}(0, 1) = \sum_{m=1}^M \omega_{tm}^3 v, \epsilon_{2t}(1, 1) = \sum_{m=1}^M \omega_{tm}^4 v \tag{12}$$

The entropy-information function  $H$  is defined as

$$\begin{aligned}
H = & - \sum_{m=1}^M p_m^\theta \ln(p_m^\theta) - \sum_{t=1}^T \sum_{m=1}^M \omega_{tm}^1 \ln(\omega_{tm}^1) - \sum_{t=1}^T \sum_{m=1}^M \omega_{tm}^2 \ln(\omega_{tm}^2) \\
& - \sum_{t=1}^T \sum_{m=1}^M \omega_{tm}^3 \ln(\omega_{tm}^3) - \sum_{t=1}^T \sum_{m=1}^M \omega_{tm}^4 \ln(\omega_{tm}^4)
\end{aligned} \tag{13}$$

---

<sup>9</sup>One could also set the variance of the random shocks equal to 1 and let  $\delta$  to be an unknown parameter to be estimated.

Our GME estimator is obtained from the estimated probabilities which are the solution to the following problem:

$$\max_{\{p_m^\theta, \omega_{tm}^1, \omega_{tm}^2, \omega_{tm}^3, \omega_{tm}^4\}} H \text{ subject to the corresponding constraints:} \quad (14)$$

$$\text{If } a_t^o = (0, 0)$$

$$\begin{aligned} 0 &> \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 0) + \sum_{m=1}^M \omega_{tm}^1 v_m \quad \text{if } 0 > \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^4 v_m \\ 0 &> \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^2 v_m \quad \text{if } 0 \leq \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^4 v_m \\ 0 &> \sum_{m=1}^M p_m^\theta z_m x_t + g(0, 1) + \sum_{m=1}^M \omega_{tm}^3 v_m \end{aligned} \quad (15)$$

$$\text{If } a_t^o = (0, 1)$$

$$\begin{aligned} 0 &> \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 0) + \sum_{m=1}^M \omega_{tm}^1 v_m \quad \text{if } 0 > \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^4 v_m \\ 0 &> \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^2 v_m \quad \text{if } 0 \leq \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^4 v_m \\ \sum_{m=1}^M p_m^\theta z_m x_t + g(0, 1) + \sum_{m=1}^M \omega_{tm}^3 v_m &\geq 0 \end{aligned} \quad (16)$$

$$\text{If } a_t^o = (1, 0)$$

$$\sum_{m=1}^M p_m^\theta z_m x_t + g(1, 0) + \sum_{m=1}^M \omega_{tm}^1 v_m \geq 0 \quad (17)$$

$$0 > \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^4 v_m$$

$$\text{If } a_t^o = (1, 1)$$

$$\sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^2 v_m \geq 0 \quad (18)$$

$$\sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^4 v_m \geq 0$$

and the normalization-additivity constraints:

$$\begin{aligned}\sum_{m=1}^M p_m^\theta &= 1 \\ \sum_{m=1}^M \omega_{tm}^i &= 1, \forall t \in T, \forall i \in \{1, 2, 3, 4\} \\ p_m^\theta, \omega_{tm}^1, \omega_{tm}^2, \omega_{tm}^3, \omega_{tm}^4 &> 0; \forall t \in T, m \in M\end{aligned}\tag{19}$$

Note that for each repetition of the game, there is a unique equilibrium. Thus, the constraints for each market are one of the four possible constraints listed in equations (15) to (18). Our GME estimator of the structural parameter  $\theta$  is

$$\hat{\theta}_{GME} = \sum_{m=1}^M \hat{p}_m^\theta z_m \tag{20}$$

For this simple game, constraints (15) and (16) contain logical if-then statements between endogenous variables. This programming is called disjunctive programming, which can be reformulated as an MINLP. For example, consider the logical statements:

$$y_1 < 0 \text{ if } x < 0; \quad y_2 < 0 \text{ if } x \geq 0 \tag{21}$$

By introducing a *zero* or *one* variable  $q$ , the statement (21) can be restated as:

$$x - K(1 - q) < 0; \quad x + Kq \geq 0; \quad y_1 < K(1 - q); \quad y_2 < Kq; \quad q = \{0, 1\}, \tag{22}$$

where  $K$  is a positive variable which exceeds the bound of  $|x|$ ,  $|y_1|$  and  $|y_2|$ . Such reformulations are discussed extensively in Williams (1985) and Raman and Grossmann (1991).

By introducing such a binary variable  $K$ , our GME method can be reformulated as:

$$\begin{aligned}
& \max_{\{p_m^\theta, \omega_{tm}^1, \omega_{tm}^2, \omega_{tm}^3, \omega_{tm}^4, q_t\}} H \text{ s.t.} \tag{23} \\
& \text{If } a_t^o = (0, 0) \left| \begin{aligned}
& \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^4 v_m \geq -K q_t \\
& \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^4 v_m < K(1 - q_t) \\
& \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 0) + \sum_{m=1}^M \omega_{tm}^1 v_m < K(1 - q_t) \\
& \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^2 v_m < K q_t \\
& 0 > \sum_{m=1}^M p_m^\theta z_m x_t + g(0, 1) + \sum_{m=1}^M \omega_{tm}^3 v_m
\end{aligned} \right. \\
& \text{if } a_t^o = (0, 1) \left| \begin{aligned}
& \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^4 v_m \geq -K q_t \\
& \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^4 v_m < K(1 - q_t) \\
& \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 0) + \sum_{m=1}^M \omega_{tm}^1 v_m < K(1 - q_t) \\
& \sum_{m=1}^M p_m^\theta z_m x_t + g(1, 1) + \sum_{m=1}^M \omega_{tm}^2 v_m < K q_t \\
& \sum_{m=1}^M p_m^\theta z_m x_t + g(0, 1) + \sum_{m=1}^M \omega_{tm}^3 v_m \geq 0
\end{aligned} \right. \\
& \text{and} \left| \begin{aligned}
& \text{equation (17), (18)} \\
& \text{normalization-additivity constraints equation (19)} \\
& q_t \in \{0, 1\}
\end{aligned} \right.
\end{aligned}$$

This optimization, which can be solved via MINLP techniques, yields the estimated probability for each unknown, which include the probabilities for our structural parameter  $\theta$ .



Thus the estimated  $\hat{\theta}$  can be recovered from the original reparameterization:

$$\hat{\theta}_{GME} = \sum_{m=1}^M \hat{p}_m^{\theta} z_m \quad (24)$$

Note that if the payoff function is linear in all covariates,  $x$ , then the optimization problem becomes a convex MINLP, since the objective entropy function is always concave. For the game which has additional players, actions, or stages, the estimation can be easily extended to involve more constraints.

The GME estimator proposed above essentially is a MPEC style estimator, as argued by [Su and Judd \(2008\)](#), implementing asymptotic inference methods is more complex within the MPEC framework. Computing standard errors requires the computation of the Hessian of the objective function with respect to the structural parameters  $\theta$ . Such work seems hard due to the non-smoothness of our GME estimation.

[Horowitz \(1997, 1998, 2001\)](#) explains how some non-smooth estimators can be smoothed in a way that greatly simplifies the analysis of their asymptotic distributional properties. Smoothing our GME estimator is also possible since for most cases we can reformulate the MINLP problem into a nonlinear programming (NLP) with complementarity constraints (MPCC). [Chen and Mangasarian \(1996\)](#) provides a class of smoothing functions for nonlinear and mixed complementarity problems.

Alternatively, some recently derived resampling methods can deal with such nonregular estimators. [Zeng and Lin \(2008\)](#) suggests some efficient resampling procedures for non-smooth estimators based on asymptotic expansion via empirical process arguments. [Andrews and Guggenberger \(2009\)](#) also provides some efficient Hybrid and Size-Corrected subsampling methods. We will investigate these alternative methods for inference within our GME estimation framework in future work.

As noted above, our framework can deal with a wide range of random shock structures which depart from the *i.i.d* assumption. One example is market specific shocks also

considered by [Maruyama \(2009\)](#). Consider the simple entry game in figure.2, where the random preference shock for player  $i$  in market  $t$  is specified as:

$$\epsilon_{it}(a) = \omega_{it}(a) + \eta_i \quad (25)$$

The market specific shock  $\eta_i$  and idiosyncratic preference shock  $\omega_{it}(a)$  are both independently distributed across players (entrants) and markets. We can individually estimate  $\omega_{it}(a)$  and  $\eta_i$  instead of  $\epsilon_{it}(a)$ , using an additive entropy objective function.

## 5 Monte Carlo

To demonstrate the performance of our estimator in small samples, we conducted two Monte Carlo experiments using the simple sequential entry game introduced in section 2. There are two players and each player has the following profit function:

$$u_i(x, a, \epsilon_i; \theta) = 1(a_i = 1) \{ \theta_1 x_1 + \theta_2 x_{i2} - \theta_3 x_{i3} + \epsilon_i(a) \} \quad (26)$$

We define  $x_1 \sim N(10, 1)$ ,  $x_{i2} \sim N(1, 1)$ , and  $x_{i3} = 9(N(a) - 1)$ , where  $N(a)$  is the number of entrants from an action profile  $a$ . In the first experiment, the idiosyncratic error terms,  $\epsilon_{it}(a)$ , are drawn from the standard normal distribution. In the second experiment,  $\epsilon_{it}(a)$  are drawn from the uniform distribution  $[-1, 1]$ .

As discussed previously, our model requires both scale and location normalizations, so we assume that  $\theta_3 = 1$  and the payoffs from not entering are zero. Thus our game has two unknown parameters:  $\theta_1$  and  $\theta_2$ . We generated 10000 samples of size  $t = 25, 50, 100$ , and 200 to assess the finite sample properties of our estimator. The true parameter vector values were chosen as  $\theta_1 = 1$  and  $\theta_2 = -1$ . The parameter estimates are presented in Tables [I](#) and [II](#). The empirical distributions of parameter estimates are reported in Figures

Table I: Monte Carlo Results for Normal Shocks

Parameter	Mean	Median	Standard Deviation	Mean Bias	Median Bias	MSE
$T = 25$						
$\theta_1$	1.1651	1.1350	0.1387	0.1651	0.1350	0.0465
$\theta_2$	-2.5561	-2.3814	0.8934	-1.5561	-1.3814	3.2194
$T = 50$						
$\theta_1$	1.0471	1.0388	0.0555	0.0471	0.0388	0.0052
$\theta_2$	-1.5566	-1.4607	0.4868	-0.5566	-0.4607	0.5466
$T = 100$						
$\theta_1$	0.9934	0.9924	0.0247	-0.0066	-0.0076	0.0006
$\theta_2$	-1.0449	-1.0265	0.2215	-0.0449	-0.0265	0.0511
$T = 200$						
$\theta_1$	0.9944	0.9941	0.0142	-0.0056	-0.0059	0.0002
$\theta_2$	-1.0403	-1.0344	0.1171	-0.0403	-0.0344	0.0153
True value: $\theta_1 = 1$ , $\theta_2 = -1$ ; Monte Carlo Times: 10000						

c and d.

The results are encouraging even in smaller samples sizes; the payoff parameters are estimated close to their true values, and as the sample size increases, the estimates become more precise. Parameter  $\theta_1$  is estimated more precisely. This is mainly because  $\theta_1$  has a larger influence over equilibrium than  $\theta_2$ , since  $\theta_1$  interacts with a covariate with a higher mean than  $\theta_2$  does. In an extremely small sample, the variation in  $\theta_2$  may not generate sufficient variation in equilibrium to be estimated precisely but this problem disappears as the sample size becomes larger.

The second simulation shows that our GME estimator can handle non-normal errors. The empirical distributions of the parameter estimates shown in figures 3 and 4 also show that our GME estimator is asymptotically normally distributed.

Table II: Monte Carlo Results for Uniform Shocks

Parameter	Mean	Median	Standard Deviation	Mean Bias	Median Bias	MSE
$T = 25$						
$\theta_1$	1.1660	1.1378	0.1410	0.1660	0.1378	0.0474
$\theta_2$	-2.5643	-2.4009	0.8977	-1.5643	-1.4009	3.2528
$T = 50$						
$\theta_1$	1.0461	1.0371	0.0548	0.0461	0.0371	0.0051
$\theta_2$	-1.5472	-1.4479	0.4841	-0.5472	-0.4479	0.5338
$T = 100$						
$\theta_1$	0.9937	0.9931	0.0253	-0.0063	-0.0069	0.0007
$\theta_2$	-1.0460	-1.0283	0.2238	-0.0460	-0.0283	0.0521
$T = 200$						
$\theta_1$	0.9946	0.9941	0.0141	0.0054	0.0059	0.0002
$\theta_2$	-1.0435	-1.0383	0.1149	0.0435	0.0383	0.0151
True value: $\theta_1 = 1$ , $\theta_2 = -1$ ; Monte Carlo Times: 10000						

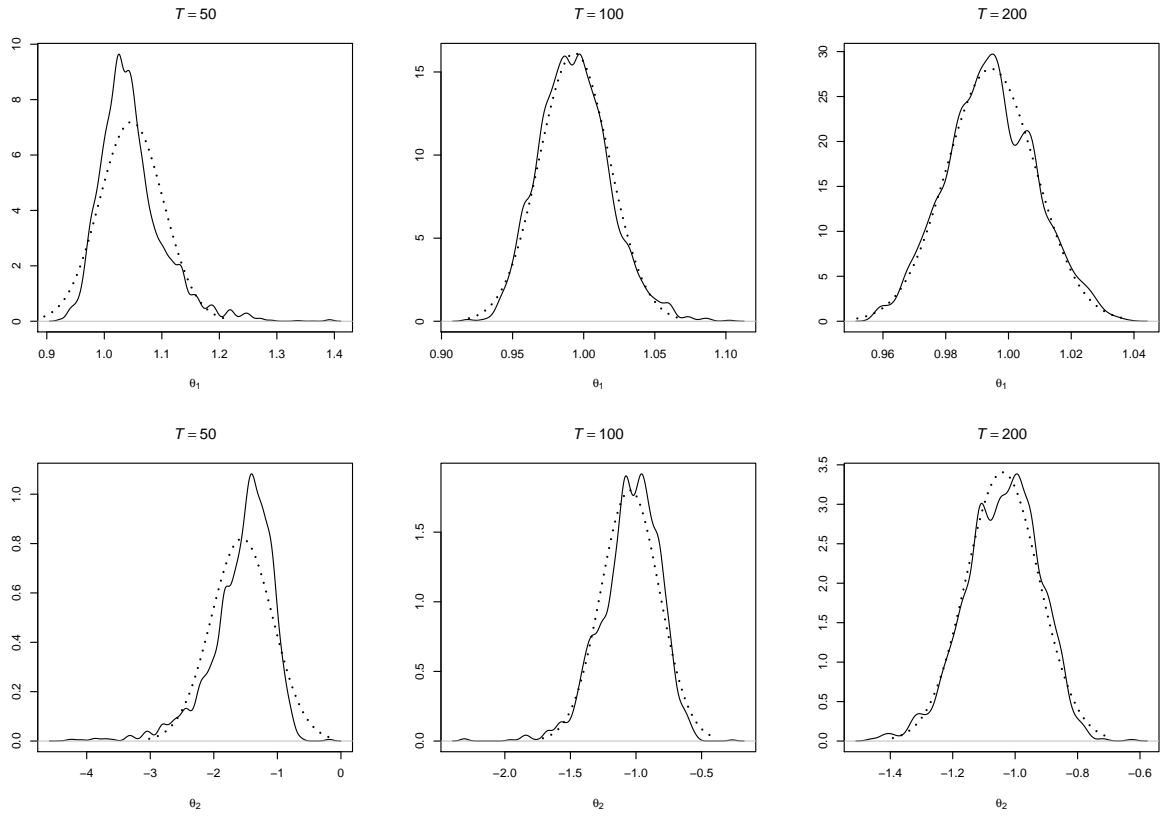


Figure c: Distribution of Estimators with Normal Shocks

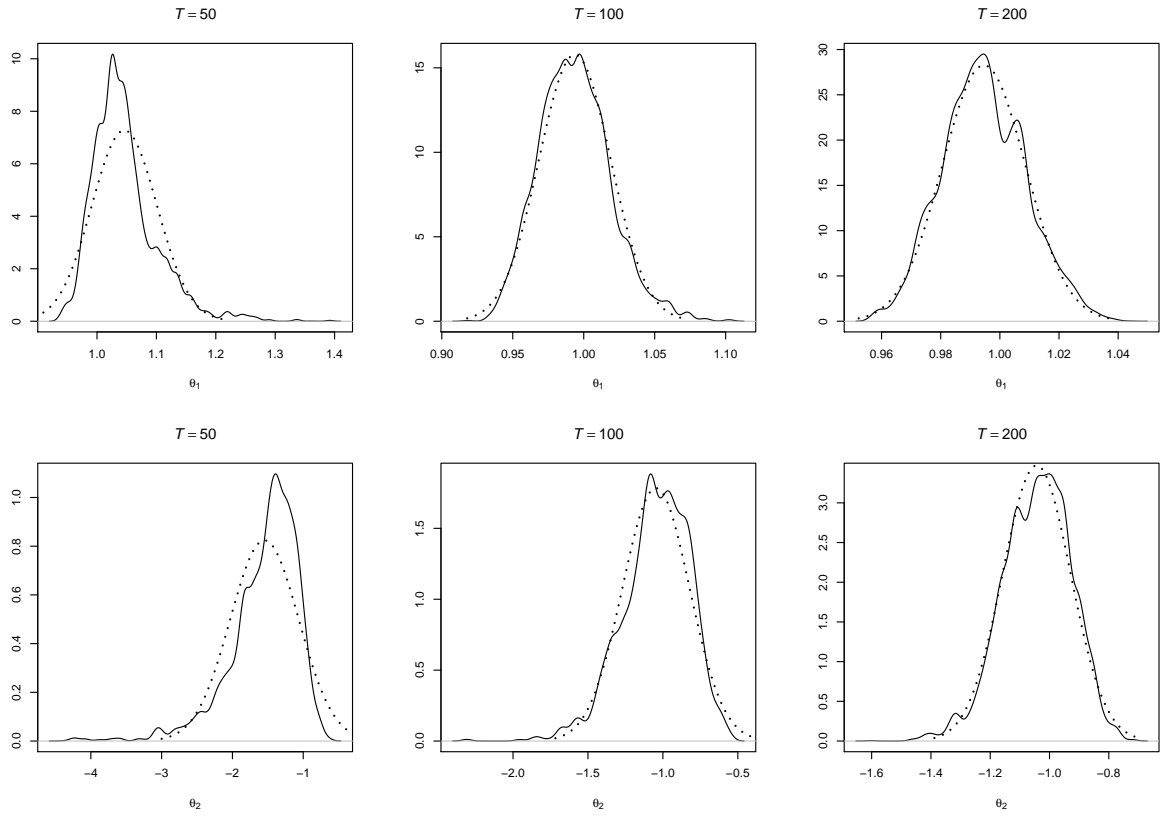


Figure d: Distribution of Estimators with Uniform Shocks

## 6 Conclusion

In this paper, we developed a data-constrained GME estimator for the discrete sequential-move game of perfect information. By directly using the data-constraints which are implied by the equilibrium conditions, we avoid the multidimensional integration which always makes such estimation intractable. Moreover, in contrast to existing estimators for general complete information games, our GME estimator requires no parametric assumption about the distribution of the random shocks. We formulate the estimation as a (convex) mixed-integer nonlinear optimization problem. The estimation can be easily implemented on optimization software with high-level interfaces such as GAMS or AMPL. The model is identified with only weak scale and location normalizations. Monte Carlo evidence demonstrates that the estimator can perform well in moderately sized samples.

## References

- AGUIRREGABIRIA, V., AND P. MIRA (2007): “Sequential Estimation of Dynamic Discrete Games,” *Econometrica*, 75, 1–53.
- ANDREWS, D. W. K., AND P. GUGGENBERGER (2009): “Hybrid and Size-Corrected Subsampling Methods,” *Econometrica*, 77(3), 721–762.
- BAJARI, P., H. HONG, J. KRAINER, AND D. NEKIPELOV (2010): “Estimating Static Models of Strategic Interactions,” *Journal of Business and Economic Statistics*.
- BAJARI, P., H. HONG, AND S. RYAN (2010): “Identification and Estimation of a Discrete Game of Complete Information,” *Econometrica*.
- BERRY, S., AND E. TAMER (2007): “Identification in Models of Oligopoly Entry,” in *Advanced in Economics and Econometrics: Theory and Application*, vol. II, chap. 2, pp. 46–85. Cambridge University Press.
- BERRY, S. T. (1992): “Estimation of a Model of Entry in the Airline Industry,” *Econometrica*, 60(4), 889–917.

- BONAMI, P., M. KILINC, AND J. LINDEROTH (2009): “Algorithms and Software for Convex Mixed Integer Nonlinear Programs,” Discussion paper, University of Wisconsin Madison.
- BRESNAHAN, T. F., AND P. C. REISS (1990): “Entry in Monopoly Markets,” *Review of Economic Studies*, 57, 531–553.
- (1991): “Empirical models of Discrete Games,” *Journal of Econometrics*, 48, 57–81.
- CAMPBELL, R. C., AND R. C. HILL (2001): “Maximum Entropy Estimation in Economic Models with Linear Inequality Restrictions,” *Working Paper*.
- CHEN, C., AND O. L. MANGASARIAN (1996): “A Class of Smoothing Functions for Nonlinear and Mixed Complementarity Problems,” *Computational Optimization and Applications*, 5(2), 97–138.
- GOLAN, A., G. JUDGE, AND J. MILLER (1996): *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. John Wiley & Sons, New York.
- GOLAN, A., G. JUDGE, AND J. PERLOFF (1997): “Estimation and Inference with Censored and Ordered Multinomial Response Data,” *Journal of Econometrics*, 79(1), 23–51.
- GOLAN, A., L. S. KARP, AND J. PERLOFF (1998): “Estimating a Mixed Strategy: United and American Airlines,” *Working Paper*.
- GOLAN, A., L. S. KARP, AND J. M. PERLOFF (2000): “Estimating Coke’s and Pepsi’s Price and Advertising Strategies,” *Journal of Business and Economic Statistics*, 18(4), 398–409.
- GOLAN, A., J. M. PERLOFF, AND E. Z. SHEN (2000): “Estimating a Demand System with Nonnegativity Constraints: Mexican Meat Demand,” *Working Paper*.
- HOROWITZ, J. L. (1997): “Bootstrap Methods in Econometrics: Theory and Numerical Performance,” in *Advances in Economics and Econometrics: Theory and Applications: Seventh World Congress*, ed. by D. M. Kreps, and K. F. Wallis, vol. III, chap. 7, pp. 188–222. Cambridge University Press.
- HOROWITZ, J. L. (1998): “Bootstrap Methods for Median Regression Models,” *Econometrica*, 66(6), 1327–1352.



- (2001): “The Bootstrap,” in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 5, pp. 3159–3228. Elsevier Science, The Netherlands.
- JAYNES, E. T. (1957a): “Information Theory and Statistical Mechanics,” *Physics Reviews*, 106, 620–630.
- (1957b): “Information Theory and Statistical Mechanics, II,” *Physics Reviews*, 108, 171–190.
- JOUNEAU-SION, F., AND O. TORRES (2006): “MMC techniques for limited dependent variables models: Implementation by the branch-and-bound algorithm,” *Journal of Econometrics*, 133(2), 479–512.
- LEE, L. (1995): “Asymptotic Bias in Simulated Maximum Likelihood Estimation of Discrete Choice Models,” *Econometric Theory*, 11(3), 437 – 483.
- MARUYAMA, S. (2009): “Estimating Sequential-move Games by a Recursive Conditioning Simulator,” *Working Paper*.
- MAZZEO, M. J. (2002): “Product Choice and Oligopoly Market Structure,” *RAND Journal of Economics*, 33(2), 221–242.
- MITTELHAMMER, R. C., AND N. S. CARDELL (1997): “The Data-Constrained GME Estimator of the GLM: Asymptotic Theory and Inference,” *Working Paper*.
- OSBORNE, M. J., AND A. RUBINSTEIN (1994): *A Course in Game Theory*. The MIT Press.
- RAMAN, R., AND I. E. GROSSMANN (1991): “Relationship Between MILP Modelling and Logical Inference for Chemical Process Synthesis,” *Computers and Chemical Engineering*, 15(2), 73–84.
- SCHMIDT-DENGLER, P. (2006): “The Timing of New Technology Adoption: The Case of MRI,” *Working Paper*.
- SHANNON, C. E. (1948): “A Mathematical Theory of Communication,” *Bell System Technical Journal*, 27, 379–423.
- SU, C.-L., AND K. L. JUDD (2008): “Constrained Optimization Approaches to Estimation of Structural Models,” *Working Paper*.

- WILLIAMS, H. P. (1985): *Model Building in Mathematical Programming*. John Wiley & Sons, New York.
- ZENG, AND LIN (2008): “Efficient Resampling Methods for Nonsmooth Estimating Functions,” *Biostatistics*, 9(2), 355–363.